

Pathwise estimates for an effective dynamics

F. Legoll^{1,4}, T. Lelièvre^{2,4} and S. Olla³

¹ Laboratoire Navier, École des Ponts ParisTech, Université Paris-Est,
6 et 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2, France

`legoll@lami.enpc.fr`

² CERMICS, École des Ponts ParisTech, Université Paris-Est,
6 et 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2, France

`lelievre@cermics.enpc.fr`

³ CEREMADE, UMR-CNRS 7534, Université de Paris Dauphine, PSL Research University
Place du Maréchal De Lattre De Tassigny, 75775 Paris Cedex 16, France

`olla@ceremade.dauphine.fr`

⁴ INRIA Paris, MATHERIALS project-team,
2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12, France

May 10, 2016

Abstract

Starting from the overdamped Langevin dynamics in \mathbb{R}^n ,

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\beta^{-1}}dW_t,$$

we consider a scalar Markov process ξ_t which approximates the dynamics of the first component X_t^1 . In the previous work [8], the fact that $(\xi_t)_{t \geq 0}$ is a good approximation of $(X_t^1)_{t \geq 0}$ is proven in terms of time marginals, under assumptions quantifying the timescale separation between the first component and the other components of X_t . Here, we prove an upper bound on the trajectorial error $\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^1 - \xi_t| \right)$, for any $T > 0$, under a similar set of assumptions. We also show that the technique of proof can be used to obtain quantitative averaging results.

1 Introduction

Coarse-graining techniques are fundamental tools in computational statistical physics problems. They are very important for modelling questions (to get some insight on a complicated high-dimensional problem, by reducing it to a low-dimensional model) and for numerical algorithms, which very often use coarse-grained descriptions as predictors to speed up the computations.

In this work, we are interested in getting a low-dimensional Markov dynamics on a few degrees of freedom starting from a high-dimensional Markov dynamics. More specifically, we consider a stochastic process $(X_t)_{t \geq 0}$ on \mathbb{R}^n which follows the overdamped Langevin dynamics:

$$dX_t = -\nabla V(X_t) dt + \sqrt{2\beta^{-1}} dW_t, \quad (1)$$

where β^{-1} is proportional to the temperature, W_t is a standard n -dimensional Brownian motion and $V : \mathbb{R}^n \rightarrow \mathbb{R}$ is a smooth function. This dynamics is often used in molecular dynamics simulation. Under suitable assumptions on V , this dynamics is ergodic with respect to the Boltzmann-Gibbs measure

$$d\mu = \psi(x) dx$$

with

$$\psi(x) = Z^{-1} \exp(-\beta V(x)), \quad Z = \int_{\mathbb{R}^n} \exp(-\beta V(x)) dx,$$

where Z is assumed to be finite. Suppose that we are not interested in the dynamics of $(X_t)_{t \geq 0}$, but only in the dynamics of its first component, $(X_t^1)_{t \geq 0}$ (see Remark 1 below for a discussion of more general cases). In view of (1), we have

$$dX_t^1 = -\partial_1 V(X_t) dt + \sqrt{2\beta^{-1}} dW_t^1. \quad (2)$$

In general, this dynamics is not closed in $(X_t^1)_{t \geq 0}$, as the right-hand side depends on X_t and not only on X_t^1 .

To obtain a closed (Markov) dynamics, a natural idea inspired by [5] is to replace the drift term in (2) by its expectation with respect to μ , conditionally to the value of X_t^1 . We refer to [8, 9] for a motivation using probabilistic arguments, and to [2, 14] for other derivations, based in particular on the so-called Mori-Zwanzig projection operator approach [3]. Roughly speaking, such an approximation requires the first component X_t^1 to “move slowly” compared to the other components (X_t^2, \dots, X_t^n) , so that these components

reach equilibrium before X_t^1 has moved. The aim of the analysis performed here and in the previous works [8, 9] is to give a precise mathematical content to this intuition.

For any $\xi \in \mathbb{R}$, we hence introduce

$$b(\xi) = \mathbb{E}_\mu \left(\partial_1 V(X) \mid X^1 = \xi \right) = \int_{\mathbb{R}^{n-1}} \partial_1 V(\xi, x_2^n) \psi^\xi(x_2^n) dx_2^n \quad (3)$$

with

$$\psi^\xi(x_2^n) = \frac{\psi(\xi, x_2^n)}{\int_{\mathbb{R}^{n-1}} \psi(\xi, x_2^n) dx_2^n}, \quad (4)$$

where we have used the notation

$$x_2^n = (x^2, \dots, x^n), \quad dx_2^n = dx^2 \dots dx^n.$$

Note that $\psi^\xi(x_2^n) dx_2^n$ is a conditional probability measure, namely the probability measure μ conditioned to $X^1 = \xi$. In the following, we also need the marginal probability density $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ of ψ along the first coordinate:

$$\varphi(\xi) = \int_{\mathbb{R}^{n-1}} \psi(\xi, x_2^n) dx_2^n. \quad (5)$$

The function b (or, depending on the authors, $-b$) is the so-called *mean force* associated to the measure μ and the *reaction coordinate* $(x^1, x_2^n) \mapsto x^1$, see e.g. [12]. It is the derivative of the so-called *free energy* $F(\xi) = -\beta^{-1} \ln \varphi(\xi)$ associated to μ and the reaction coordinate $(x^1, x_2^n) \mapsto x^1$: $b(\xi) = F'(\xi)$.

Replacing the drift term in (2) by its conditional expectation, we introduce the following dynamics, which hopefully is a good approximation of (2):

$$\begin{cases} d\xi_t = -b(\xi_t) dt + \sqrt{2\beta^{-1}} dW_t^1, \\ \xi_0 = X_0^1. \end{cases} \quad (6)$$

This is a closed dynamics in ξ_t , and $(\xi_t)_{t \geq 0}$ is a Markov process. The mathematical question is now to estimate, in some sense to be made precise, the distance between ξ_t solution to (6) and X_t^1 which satisfies (2)–(1). By construction, the effective dynamics (6) has the correct stationary state: it is ergodic with respect to $\varphi(\xi) d\xi$, which is precisely the law of X_t^1 in the long-time limit. The question we address here concerns the correctness of the

dynamics. This is motivated in particular by current practices in the field of molecular dynamics, where practitioners derive from (6) some transition times by looking at free energy differences and using the Eyring-Kramers law (see for example [6] for a review).

As a first step, in [8], estimates on the distance (in total variation norm) between the law at time t of X_t^1 and the law at time t of ξ_t have been obtained. These are therefore estimates on the distance between the marginals in time. More precisely, under the two following assumptions:

[A1] The conditional probability measures $\psi^\xi(x_2^n) dx_2^n$ satisfy a Logarithmic Sobolev Inequality with a constant ρ which does not depend on ξ ,

[A2] The so-called coupling constant κ_∞ is finite:

$$\kappa_\infty = \|\partial_1 \widehat{\nabla} V\|_{L^\infty(\mathbb{R}^n)} < \infty,$$

where $\widehat{\nabla} V = (\partial_2 V, \dots, \partial_n V)$,

we have shown in [8] (see also [9] for a simple case) that, for all times $t \geq 0$,

$$H\left(\mathcal{L}(X_t^1) \mid \mathcal{L}(\xi_t)\right) \leq \frac{\beta^2 \kappa_\infty^2}{4\rho^2} \left[H\left(\mathcal{L}(X_0) \mid \mu\right) - H\left(\mathcal{L}(X_t) \mid \mu\right) \right]. \quad (7)$$

Here, $H\left(\mathcal{L}(X_t^1) \mid \mathcal{L}(\xi_t)\right) = \int_{\mathbb{R}} \ln \left(\frac{d\mathcal{L}(X_t^1)}{d\mathcal{L}(\xi_t)} \right) d\mathcal{L}(X_t^1)$ denotes the relative entropy of the law of X_t^1 with respect to the law of ξ_t , which is for example an upper bound on the square distance between the two laws in total variation norm. Provided that ρ is large, we hence see that, in terms of laws at any time t , ξ_t is an accurate approximation of X_t^1 . The assumption that ρ is large formalizes the fact that mixing with respect to the probability measures $\psi^\xi(x_2^n) dx_2^n$ is fast, and hence, as pointed out above, that the components (X_t^2, \dots, X_t^n) quickly reach equilibrium.

As a side remark, let us mention that the set of assumptions [A1]–[A2] appears to be very useful to analyze metastable processes and coarse-graining techniques in many contexts (see e.g. [4, 10, 11]).

In this work, we go further and obtain estimates on some distance between the trajectories $(X_t^1)_{t \geq 0}$ and $(\xi_t)_{t \geq 0}$ (and not only the laws of X_t^1 and ξ_t at any time t), under assumptions very similar to Assumptions [A1]–[A2] (see Proposition 3 below). The main difference is that we obtain a result over

finite time intervals, whereas (7) is a uniform in time estimate. Getting some trajectorial estimates is crucial since, in molecular dynamics, many quantities of interest are indeed trajectorial ones (such as autocorrelation in time of some observables, for example).

The article is organized as follows. In Section 2, we introduce some notation, the assumptions under which we work, and our main result (Proposition 3). This result is based on three ingredients:

- first, an estimate on $\mathbb{E}_\mu(f^2)$ where $f = b - \partial_1 V$, which is a direct consequence of two assumptions similar to the assumptions [A1] and [A2] above, see Section 3;
- second, the introduction of a Poisson equation and the use of an argument due to T. Lyons and T. Zhang in [13] to get an estimate on $\mathbb{E}_\mu \left(\sup_{0 \leq t \leq T} \left| \int_0^t f(X_s) ds \right|^2 \right)$, see Section 4;
- third, a Gronwall type argument to deduce from this estimate a bound on $\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^1 - \xi_t| \right)$. This argument requires some Lipschitz type assumptions on b , see Section 5.

Section 6 collects all these results to conclude the proof of Proposition 3. Finally, as an application of the mathematical techniques used to prove Proposition 3, we provide in Section 7 a quantitative averaging result, which is to the best of our knowledge new since it does not require the effective drift b to be Lipschitz.

Remark 1. *We consider here for simplicity the case when the degree of freedom of interest is one of the cartesian coordinate of X_t . This could be generalized in two directions. First, one could consider a more general function of X_t , say $\theta(X_t)$, where $\theta : \mathbb{R}^n \rightarrow \mathbb{R}$. This is actually the setting of the previous work [8]. One could also consider higher dimensional settings, where θ takes its values in \mathbb{R}^d with $d \geq 2$. We do not pursue along these directions here, in order to keep the presentation simple.*

2 Notation, assumptions and main result

2.1 Notation

Let us introduce the operator L , defined by: for any function $v : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\begin{aligned} Lv &= -\nabla V \cdot \nabla v + \beta^{-1} \Delta v \\ &= -\sum_{i=1}^n \partial_i V \partial_i v + \beta^{-1} \sum_{i=1}^n \partial_{ii} v. \end{aligned} \quad (8)$$

We also need the family of operators L^ξ indexed by $\xi \in \mathbb{R}$ and defined by: for any function $v : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\begin{aligned} (L^\xi v)(\xi, x_2^n) &= -\widehat{\nabla} V(\xi, x_2^n) \cdot \widehat{\nabla} v(\xi, x_2^n) + \beta^{-1} \widehat{\Delta} v(\xi, x_2^n) \\ &= -\sum_{i=2}^n \partial_i V(\xi, x_2^n) \partial_i v(\xi, x_2^n) + \beta^{-1} \sum_{i=2}^n \partial_{ii} v(\xi, x_2^n), \end{aligned} \quad (9)$$

where we used the notation

$$\widehat{\nabla} v = (\partial_2 v, \dots, \partial_n v) \quad \text{and} \quad \widehat{\Delta} v = \sum_{i=2}^n \partial_{ii} v.$$

Note that the sums in (9) start at $i = 2$, in contrast with those in (8).

The functional space

$$L^2(\psi) = \left\{ v : \mathbb{R}^n \rightarrow \mathbb{R}, \ v \in L^1_{\text{loc}}(\mathbb{R}^n) \text{ and } \int_{\mathbb{R}^n} v^2 \psi < \infty \right\}$$

plays a crucial role in the following. It is an Hilbert space for the scalar product: for u and v in $L^2(\psi)$,

$$\langle u, v \rangle_\psi = \int_{\mathbb{R}^n} u v \psi.$$

Likewise, we will use the space $L^2(\psi^\xi)$, defined over functions $v : x_2^n \in \mathbb{R}^{n-1} \mapsto v(x_2^n) \in \mathbb{R}$.

For a given operator O , we denote by O^* its adjoint with respect to the scalar product in $L^2(\psi)$: for any smooth test functions u and v ,

$$\langle O^* u, v \rangle_\psi = \langle u, Ov \rangle_\psi.$$

It is standard to check that L is a symmetric operator in $L^2(\psi)$ (which is equivalent to the reversibility of the process X_t with respect to the equilibrium measure μ): for any smooth test functions u and v ,

$$\langle Lu, v \rangle_\psi = \langle u, Lv \rangle_\psi = -\beta^{-1} \int_{\mathbb{R}^n} \nabla u \cdot \nabla v \, \psi.$$

We thus have $L^\star = L$, and $L = -\beta^{-1} \nabla^\star \nabla$.

2.2 Assumptions

In the sequel, we work under the three following assumptions.

First, we assume that, for any ξ , the conditional probability measures $\psi^\xi(x_2^n) dx_2^n$ defined by (4) satisfy a Poincaré inequality for a constant ρ independent of ξ : there exists $\rho > 0$ such that, for any ξ and any function $v \in H^1(\psi^\xi)$, it holds:

$$\begin{aligned} \int_{\mathbb{R}^{n-1}} \left(v(x_2^n) - \int_{\mathbb{R}^{n-1}} v(x_2^n) \psi^\xi(x_2^n) dx_2^n \right)^2 \psi^\xi(x_2^n) dx_2^n \\ \leq \frac{1}{\rho} \int_{\mathbb{R}^{n-1}} \left| \widehat{\nabla} v(x_2^n) \right|^2 \psi^\xi(x_2^n) dx_2^n, \end{aligned} \quad (10)$$

where the functional space $H^1(\psi^\xi)$ is defined by

$$H^1(\psi^\xi) = \left\{ v : \mathbb{R}^{n-1} \rightarrow \mathbb{R}, \int_{\mathbb{R}^{n-1}} \left(v^2 + |\widehat{\nabla} v|^2 \right) \psi^\xi < \infty \right\}. \quad (11)$$

Note that, by integration by parts, we have

$$\int_{\mathbb{R}^{n-1}} (-L^\xi \phi)(x_2^n) \phi(x_2^n) \psi^\xi(x_2^n) dx_2^n = \beta^{-1} \int_{\mathbb{R}^{n-1}} \left| \widehat{\nabla} \phi(x_2^n) \right|^2 \psi^\xi(x_2^n) dx_2^n$$

for any $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^{n-1})$. The assumption (10) is thus a spectral gap assumption on the operator $-L^\xi$. A Poincaré inequality holds on a probability measure $\exp(-\beta W(x)) dx$ under relatively mild assumption on W . Basically, if W is smooth and grows at least linearly at infinity, then $\exp(-\beta W(x)) dx$ satisfies a Poincaré inequality (see for example [1]). In particular, if W is α -convex, then the Poincaré inequality is satisfied with the constant $\alpha/2$.

Second, we assume that the cross derivative $\widehat{\nabla} \partial_1 V$ is in $L^2(\psi)$:

$$\kappa^2 := \int_{\mathbb{R}^n} \left| \widehat{\nabla} \partial_1 V(x) \right|^2 \psi(x) dx < \infty. \quad (12)$$

The two above assumptions (10) and (12) are very similar to (and actually weaker than) the assumptions [A1] and [A2] mentioned in the introduction and which have been used in [8] to study the correctness of the effective dynamics in terms of time marginals.

Third, we assume that the function b defined by (3) is one-sided Lipschitz on \mathbb{R} : there exists $L_b > 0$ such that

$$\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \quad (b(y) - b(x)) (x - y) \leq L_b (x - y)^2. \quad (13)$$

If b is differentiable, this is equivalent to $-b'(x) \leq L_b$ for any $x \in \mathbb{R}$.

In addition, for any $x > 0$, we introduce

$$\alpha(x) = \sup_{s \in [-x, x]} |b'(s)| \quad (14)$$

and assume that

$$\begin{aligned} C_\alpha(\beta) &= \mathbb{E} \left[\left(\alpha(|X^1|) \right)^2 \right] = \mathbb{E} \left[\left(\sup_{s \in [-|X^1|, |X^1|]} |b'(s)| \right)^2 \right] \\ &= \int_{\mathbb{R}} \left(\sup_{s \in [-|\xi|, |\xi|]} |b'(s)| \right)^2 \varphi(\xi) d\xi < \infty \end{aligned} \quad (15)$$

where φ (defined by (5)) is the marginal probability density along the first coordinate X^1 . The quantity C_α depends on β as φ and b depend on β . Note that, if we think of V as having polynomial growth, we see that α also has polynomial growth. In this case, the assumption (15) is hence fulfilled. The assumption (15) is further discussed in Remarks 13 and 14 below.

We will also sometimes consider the assumption

$$C_{\alpha,p}(\beta) = \int_{\mathbb{R}} \left(\sup_{s \in [-|\xi|, |\xi|]} |b'(s)| \right)^{2p/(2-p)} \varphi(\xi) d\xi < \infty \quad (16)$$

for some $1 \leq p < 2$, which is stronger than (15). Note that (15) corresponds to the case $p = 1$.

Roughly speaking, the assumptions (13) and (15) will be used below to show that if $x(t)$ and $y(t)$ are solutions to $\dot{x} = -b(x)$ and $\dot{y} = -b(y) + \dot{e}$ (with $x(0) = y(0)$), then $\|x - y\|_{L^\infty(0,T)}$ is small if $\|e\|_{L^\infty(0,T)}$ is small (see Lemma 12 below).

Remark 2. *The assumption (13) is satisfied if b is Lipschitz on bounded domains and increasing at infinity, which corresponds to a case when the associated free energy F is convex at infinity and smooth.*

2.3 Main result

Our main result is the following.

Proposition 3. *Assume that (10), (12), (13) and (15) hold, and that the system starts at equilibrium:*

$$X_0 \sim \mu. \quad (17)$$

Consider $(X_t)_{0 \leq t \leq T}$ solution to (1) and $(\xi_t)_{0 \leq t \leq T}$ solution to (6) over a bounded time interval $[0, T]$. Then, there exists a constant C , that is independent of ρ and κ , and that only depends on T , $C_\alpha(\beta)$ and L_b , such that

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^1 - \xi_t| \right) \leq C \sqrt{\beta} \frac{\kappa}{\rho}. \quad (18)$$

Note that the constant C in (18) only depends on β through its dependency to $C_\alpha(\beta)$.

Remark 4. *The proof of (18) also shows that, if we replace the assumption (15) by the stronger assumption (16) for some $1 \leq p < 2$, then there exists a constant C , that is independent of ρ and κ , and that only depends on p , T , $C_{\alpha,p}(\beta)$ and L_b , such that*

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^1 - \xi_t|^p \right) \leq C \left(\sqrt{\beta} \frac{\kappa}{\rho} \right)^p. \quad (19)$$

See Remark 15 below.

Remark 5. *If $\partial_1 V$ is independent of x_2^n , then the dynamics (2) is actually closed in X^1 and we expect the effective dynamics (6) to be exact. This is indeed the case: if $\partial_1 V$ is independent of x_2^n , then we see from (12) that $\kappa = 0$, and (18) implies that the effective dynamics is exact.*

Before going into the details in the next sections, let us sketch the proof of Proposition 3. We introduce

$$f(x) = b(x^1) - \partial_1 V(x) \quad (20)$$

and recast (2) in the form

$$dX_t^1 = -b(X_t^1) dt + f(X_t) dt + \sqrt{2\beta^{-1}} dW_t^1.$$

Using the definition (3) of b , one gets: for any $\xi \in \mathbb{R}$,

$$\int_{\mathbb{R}^{n-1}} f(\xi, x_2^n) \psi^\xi(x_2^n) dx_2^n = 0.$$

Hence, under some adequate assumptions, for any ξ , there exists a unique function $x_2^n \mapsto u(\xi, x_2^n)$ solution to the following Poisson problem:

$$L^\xi u = f(\xi, \cdot) \quad \text{with} \quad \int_{\mathbb{R}^{n-1}} u(\xi, x_2^n) \psi^\xi(x_2^n) dx_2^n = 0. \quad (21)$$

A precise result will be stated below (see Lemma 9). On the one hand, we thus obtain, after integration in time, that

$$X_t^1 - X_0^1 = - \int_0^t b(X_s^1) ds + \sqrt{2\beta^{-1}} W_t^1 + \int_0^t L^{X_s^1} u(X_s) ds. \quad (22)$$

On the other hand, integrating the stochastic differential equation (6) between times 0 and t , we get

$$\xi_t - X_0^1 = - \int_0^t b(\xi_s) ds + \sqrt{2\beta^{-1}} W_t^1. \quad (23)$$

We deduce from (22) and (23) that

$$X_t^1 - \xi_t = \int_0^t (b(\xi_s) - b(X_s^1)) ds + \int_0^t L^{X_s^1} u(X_s) ds. \quad (24)$$

We then see that we need to bound $\int_0^t L^{X_s^1} u(X_s) ds = \int_0^t f(X_s) ds$ in the right-hand side of (24) in order to estimate the distance between X_t^1 and ξ_t .

The proof is then based on two main arguments. First, we estimate the term $\int_0^t L^{X_s^1} u(X_s) ds$ using a result due to T. Lyons and T. Zhang in [13] together with an estimate on the solution u to the Poisson problem (21), see Section 4. This estimate relies on the two assumptions (10) and (12).

Second, a Gronwall-type result is proved in Section 5 to obtain an upper bound on $\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^1 - \xi_t| \right)$ in terms of $\sqrt{\mathbb{E} \left(\sup_{0 \leq t \leq T} \left| \int_0^t L^{X_s^1} u(X_s) ds \right|^2 \right)}$. This result (of interest by its own) relies on the one-sided Lipschitz assumption (13) as well as on the integrability assumption (15).

We eventually point out that the stationarity assumption (17) in Proposition 3 can be relaxed using a standard argument based on a conditional expectation with respect to the initial condition, as stated in the following corollary:

Corollary 6. *Assume that (10), (12), (13) and (15) hold. Consider $(X_t)_{0 \leq t \leq T}$ solution to (1) and $(\xi_t)_{0 \leq t \leq T}$ solution to (6) over a bounded time interval $[0, T]$, with the initial condition X_0 distributed according to a measure $\psi_0(x) dx$ such that*

$$m = \left\| \frac{\psi_0}{\psi} \right\|_{L^\infty(\mathbb{R}^n)} < \infty. \quad (25)$$

Then we have

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^1 - \xi_t| \right) \leq C m \sqrt{\beta} \frac{\kappa}{\rho}, \quad (26)$$

where C is the constant of the estimate (18).

Proof. Let us introduce

$$h(x_0) = \mathbb{E}^{x_0} \left(\sup_{0 \leq t \leq T} |X_t^1 - \xi_t| \right)$$

where \mathbb{E}^{x_0} is the expectation conditionally to the fact that the initial condition of (1) is deterministic: $X_0 = x_0 \in \mathbb{R}^n$. The result of Proposition 3 is that

$$\int_{\mathbb{R}^n} h(x) \psi(x) dx \leq C \sqrt{\beta} \frac{\kappa}{\rho}.$$

Let us now consider $(X_t^1)_{t \geq 0}$ solution to (1) with initial condition distributed according to $\psi_0(x) dx$. We have, using (25),

$$\begin{aligned} \mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^1 - \xi_t| \right) &= \int_{\mathbb{R}^n} h(x) \psi_0(x) dx = \int_{\mathbb{R}^n} h(x) \psi(x) \frac{\psi_0(x)}{\psi(x)} dx \\ &\leq \left\| \frac{\psi_0}{\psi} \right\|_{L^\infty(\mathbb{R}^n)} \int_{\mathbb{R}^n} h(x) \psi(x) dx \leq C m \sqrt{\beta} \frac{\kappa}{\rho}. \end{aligned}$$

This concludes the proof of Corollary 6. \square

A similar corollary can be stated if we assume (16) rather than (15), under the assumption that $\left\| \frac{\psi_0}{\psi} \right\|_{L^q(\psi)} < \infty$ for some q (see Remark 4).

3 An estimate on f and a preliminary result

3.1 Estimate on f

A direct consequence of the two assumptions (10) and (12) is an estimate on the function f defined by (20):

$$f(x) = b(x^1) - \partial_1 V(x) = \int_{\mathbb{R}^{n-1}} \partial_1 V(x^1, x_2^n) \psi^{x^1}(x_2^n) dx_2^n - \partial_1 V(x).$$

Lemma 7. *Consider f defined by (20) and assume that the conditional probability measures $\psi^\xi(x_2^n) dx_2^n$ satisfy the Poincaré inequalities (10), and that the cross derivative $\partial_1 \widehat{\nabla} V$ satisfies (12). Then we have*

$$\int_{\mathbb{R}^n} f^2 \psi \leq \frac{\kappa^2}{\rho}. \quad (27)$$

Proof. Notice that for any $\xi \in \mathbb{R}$, $\int_{\mathbb{R}^{n-1}} f(\xi, x_2^n) \psi^\xi(x_2^n) dx_2^n = 0$. Thus, using (10), we get, for any fixed $\xi \in \mathbb{R}$,

$$\int_{\mathbb{R}^{n-1}} (f(\xi, x_2^n))^2 \psi^\xi(x_2^n) dx_2^n \leq \rho^{-1} \int_{\mathbb{R}^{n-1}} \left| \widehat{\nabla} \partial_1 V(\xi, x_2^n) \right|^2 \psi^\xi(x_2^n) dx_2^n.$$

By multiplying by $\varphi(\xi)$, integrating over $\xi \in \mathbb{R}$ and using (12), we obtain (27). \square

Notice that, as a corollary of (27), since

$$\int_{\mathbb{R}} \left[\int_{\mathbb{R}^{n-1}} (f(\xi, x_2^n))^2 \psi^\xi(x_2^n) dx_2^n \right] \varphi(\xi) d\xi = \int_{\mathbb{R}^n} f^2 \psi < \infty$$

and $\varphi > 0$, we have that, for almost all $\xi \in \mathbb{R}$,

$$\int_{\mathbb{R}^{n-1}} f^2(\xi, x_2^n) \psi^\xi(x_2^n) dx_2^n < \infty. \quad (28)$$

3.2 A simple consequence of Lemma 7

Before proving our main result Proposition 3, we first state a preliminary result, which is weaker but also much more simple to prove than Proposition 3. This result already highlights the importance of the assumptions (10), (12) and (13).

Lemma 8. *Assume that (10), (12), (13) and (17) hold. Consider $(X_t)_{0 \leq t \leq T}$ solution to (1) and $(\xi_t)_{0 \leq t \leq T}$ solution to (6) over a bounded time interval $[0, T]$. Then, there exists a constant C , that depends on T and L_b , but is independent of β , ρ and κ , such that*

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} (X_t^1 - \xi_t)^2 \right) \leq C \frac{\kappa^2}{\rho}. \quad (29)$$

Note that we do not need the assumption (15) here. The above result obviously implies that

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^1 - \xi_t| \right) \leq C \frac{\kappa}{\sqrt{\rho}},$$

a result which is weaker than Proposition 3 as we think of ρ as being large. The stationarity assumption (17) can be weakened in a similar way as in Corollary 6 above by using the same conditioning argument.

Proof. From (2) and (6), and using the definition (20) of f , we have

$$d(X_t^1 - \xi_t) = (b(\xi_t) - b(X_t^1)) dt + f(X_t) dt.$$

We deduce from an Itô's computation that

$$\begin{aligned} \frac{1}{2} (X_t^1 - \xi_t)^2 &= \int_0^t (X_s^1 - \xi_s) (b(\xi_s) - b(X_s^1)) ds + \int_0^t (X_s^1 - \xi_s) f(X_s) ds \\ &\leq \left(L_b + \frac{1}{2}\right) \int_0^t (\xi_s - X_s^1)^2 ds + \frac{1}{2} \int_0^t (f(X_s))^2 ds, \end{aligned}$$

where we have used (13) and a Young inequality in the last line. Setting $\phi(t) = \frac{1}{2} \int_0^t (X_s^1 - \xi_s)^2 ds$ and $M = \frac{1}{2} \int_0^T (f(X_s))^2 ds$, we thus see that

$$\phi'(t) \leq M + (2L_b + 1)\phi(t). \quad (30)$$

Using Gronwall lemma and the fact that $\phi(0) = 0$, we deduce that, for any $t \in [0, T]$,

$$\phi(t) \leq \frac{M}{2L_b + 1} (e^{(2L_b+1)t} - 1).$$

Hence, using (30), we obtain

$$\frac{1}{2} (X_t^1 - \xi_t)^2 = \phi'(t) \leq e^{(2L_b+1)t} M \leq \frac{1}{2} e^{(2L_b+1)T} \int_0^T (f(X_s))^2 ds,$$

where the right-hand side is independent of t . Taking the supremum over $t \in [0, T]$ and taking expectations, we deduce that

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} (X_t^1 - \xi_t)^2 \right) \leq e^{(2L_b+1)T} \mathbb{E} \left[\int_0^T (f(X_s))^2 ds \right].$$

Now, one can use Lemma 7 above to control the right-hand side. Indeed, since X_0 (and thus X_t at any time t) is distributed according to the equilibrium measure (see (17)), we obtain, using (27), that

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} (X_t^1 - \xi_t)^2 \right) \leq e^{(2L_b+1)T} T \int_{\mathbb{R}^n} f^2 \psi \leq T e^{(2L_b+1)T} \frac{\kappa^2}{\rho}.$$

This proves the claimed bound (29) and concludes the proof of Lemma 8. \square

4 Estimate on the term $\int_0^t L^{X_s^1} u(X_s) ds$ in (24)

The aim of this section is to get an estimate in terms of κ and ρ on the last term $\int_0^t L^{X_s^1} u(X_s) ds = \int_0^t f(X_s) ds$ in (24) (where, we recall, u is the solution to the Poisson problem (21)) using the estimate on f of Lemma 7, and assuming that X_0 is distributed according to the equilibrium measure μ . As explained in Section 4.2, it is enough to estimate $\int_{\mathbb{R}^n} |\widehat{\nabla} u|^2 \psi$. The well-posedness of (21) and a bound on $\int_{\mathbb{R}^n} |\widehat{\nabla} u|^2 \psi$ is shown in Section 4.1. Finally, Section 4.3 collects the results of Sections 4.1 and 4.2 to get an estimate on $\int_0^t L^{X_s^1} u(X_s) ds$.

In all this section, only the two assumptions (10) and (12) are needed.

4.1 The Poisson problem

Let us first state an existence and uniqueness result for the Poisson problem (21) introduced above, as well as an estimate on its solution.

Lemma 9. *Assume that (10) and (12) hold, and consider the function f defined by (20). Then, for any $\xi \in \mathbb{R}$, there exists a unique function $x_2^n \mapsto u(\xi, x_2^n)$ in $H^1(\psi^\xi)$ such that*

$$L^\xi u = f(\xi, \cdot) \quad \text{with} \quad \int_{\mathbb{R}^{n-1}} u(\xi, x_2^n) \psi^\xi(x_2^n) dx_2^n = 0, \quad (31)$$

where the functional space $H^1(\psi^\xi)$ is defined by (11). Moreover, u is a \mathcal{C}^∞ function and satisfies the estimate

$$\int_{\mathbb{R}^{n-1}} |\widehat{\nabla} u(\xi, x_2^n)|^2 \psi^\xi(x_2^n) dx_2^n \leq \frac{\beta^2}{\rho} \int_{\mathbb{R}^{n-1}} f^2(\xi, x_2^n) \psi^\xi(x_2^n) dx_2^n. \quad (32)$$

In addition, we have

$$\int_{\mathbb{R}^n} |\widehat{\nabla} u|^2 \psi \leq \beta^2 \frac{\kappa^2}{\rho^2}. \quad (33)$$

Proof. Let us introduce the functional space

$$H_m^1(\psi^\xi) = \left\{ v \in H^1(\psi^\xi), \quad \int_{\mathbb{R}^{n-1}} v(x_2^n) \psi^\xi(x_2^n) dx_2^n = 0 \right\}.$$

A variational formulation of (31) is the following: find $u(\xi, \cdot) \in H_m^1(\psi^\xi)$ such that, for all $v \in H_m^1(\psi^\xi)$,

$$-\beta^{-1} \int_{\mathbb{R}^{n-1}} \widehat{\nabla} u(\xi, \cdot) \cdot \widehat{\nabla} v \psi^\xi = \int_{\mathbb{R}^{n-1}} f(\xi, \cdot) v \psi^\xi. \quad (34)$$

Here, we used the fact that for any smooth functions $u : \mathbb{R}^n \rightarrow \mathbb{R}$ and $v : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$,

$$\int_{\mathbb{R}^{n-1}} (L^\xi u) v \psi^\xi = -\beta^{-1} \int_{\mathbb{R}^{n-1}} \widehat{\nabla} u(\xi, \cdot) \cdot \widehat{\nabla} v \psi^\xi.$$

The variational problem (34) admits a unique solution using Lax-Milgram lemma and (10) to get the coercivity of the bilinear form. The right-hand-side in (34) is well defined in view of (28). Moreover, the solution to (34) is indeed a solution to (31) (in distributional sense, say) since $\int_{\mathbb{R}^{n-1}} f \psi^\xi = 0$.

By standard elliptic regularity results, since the functions V and f are assumed to be smooth, the function u is actually \mathcal{C}^∞ .

By taking $v = u$ in (34), we get

$$\begin{aligned} \int_{\mathbb{R}^{n-1}} |\widehat{\nabla} u|^2 \psi^\xi &= -\beta \int_{\mathbb{R}^{n-1}} f u \psi^\xi \\ &\leq \beta \left(\int_{\mathbb{R}^{n-1}} f^2 \psi^\xi \right)^{1/2} \left(\int_{\mathbb{R}^{n-1}} u^2 \psi^\xi \right)^{1/2} \\ &\leq \frac{\beta}{\sqrt{\rho}} \left(\int_{\mathbb{R}^{n-1}} f^2 \psi^\xi \right)^{1/2} \left(\int_{\mathbb{R}^{n-1}} |\widehat{\nabla} u|^2 \psi^\xi \right)^{1/2} \end{aligned}$$

where we used (10) in the last line. This yields (32). By combining (27) and (32), we get (33). This concludes the proof. \square

4.2 An estimate on square-integrable martingales

The following general result (see [13] or [7, Section 2.5, Lemma 2.4]) is useful for our proof.

Lemma 10. *Let $(X_t)_{t \geq 0}$ be the solution to (1), with its initial condition distributed according to the equilibrium measure μ (see assumption (17)). Consider a function $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $\Phi \in (C^\infty \cap L^2(\psi))^n$. Then, for any T , we have*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t \nabla^* \Phi(X_s) ds \right|^2 \right] \leq 8T\beta \|\Phi\|_{L^2(\psi)}^2, \quad (35)$$

where, we recall, ∇^* denotes the adjoint of the operator ∇ with respect to the $L^2(\psi)$ scalar product, so that

$$\nabla^* \Phi = \beta \nabla V \cdot \Phi - \operatorname{div}(\Phi).$$

We have not looked at the minimal regularity assumptions on Φ for this lemma to hold, since we use it below for a function Φ which is indeed in $(C^\infty \cap L^2(\psi))^n$. We refer to [7, Section 2.5, Lemma 2.4] for a statement under weaker regularity assumptions on Φ .

Proof. The proof falls in two steps.

Step 1. For any $\eta > 0$, consider the resolvent problem

$$\eta w_\eta - Lw_\eta = -\nabla^* \Phi, \quad (36)$$

a variational formulation of which is: find $w_\eta \in H^1(\psi)$ such that, for any test function $v \in H^1(\psi)$,

$$\eta \int_{\mathbb{R}^n} w_\eta v \psi + \beta^{-1} \int_{\mathbb{R}^n} \nabla w_\eta \cdot \nabla v \psi = - \int_{\mathbb{R}^n} \Phi \cdot \nabla v \psi.$$

Using the Lax-Milgram theorem, the above problem has a unique solution $w_\eta \in H^1(\psi)$. Furthermore, taking $v \equiv w_\eta$ as function test in the above variational formulation, we get

$$\eta \|w_\eta\|_{L^2(\psi)}^2 + \beta^{-1} \|\nabla w_\eta\|_{L^2(\psi)}^2 \leq \|\Phi\|_{L^2(\psi)} \|\nabla w_\eta\|_{L^2(\psi)},$$

which hence shows that, for any $\eta > 0$, we have

$$\|\nabla w_\eta\|_{L^2(\psi)} \leq \beta \|\Phi\|_{L^2(\psi)}. \quad (37)$$

Furthermore, we have $\sqrt{\eta} \|w_\eta\|_{L^2(\psi)} \leq \sqrt{\beta} \|\Phi\|_{L^2(\psi)}$, hence

$$\lim_{\eta \rightarrow 0} \eta \|w_\eta\|_{L^2(\psi)} = 0. \quad (38)$$

In addition, by standard elliptic regularity results, since the functions V and Φ are assumed to be smooth, the function w_η is actually \mathcal{C}^∞ .

Step 2. Now, let us consider a fixed time $T > 0$. Since w_η is smooth, we can write, by Itô's calculus, that, for any $t \in [0, T]$,

$$w_\eta(X_t) - w_\eta(X_0) = \int_0^t Lw_\eta(X_s) ds + \sqrt{2\beta^{-1}} \int_0^t \nabla w_\eta(X_s) \cdot dW_s. \quad (39)$$

Let us introduce, for $s \in [0, T]$,

$$Y_s = X_{T-s}.$$

Since X_0 is distributed according to the equilibrium measure μ and $(X_t)_{t \geq 0}$ is reversible with respect to μ , $(Y_s)_{0 \leq s \leq T}$ has the same law as $(X_s)_{0 \leq s \leq T}$. We thus can write

$$dY_s = -\nabla V(Y_s) ds + \sqrt{2\beta^{-1}} d\bar{W}_s$$

with $Y_0 = X_T$ and where $(\bar{W}_s)_{0 \leq s \leq T}$ is a Brownian motion. Similarly to (39), we have, for any $t \in [0, T]$,

$$w_\eta(Y_T) - w_\eta(Y_{T-t}) = \int_{T-t}^T Lw_\eta(Y_s) ds + \sqrt{2\beta^{-1}} \int_{T-t}^T \nabla w_\eta(Y_s) \cdot d\bar{W}_s. \quad (40)$$

Setting

$$M_t = \int_0^t \nabla w_\eta(X_s) \cdot dW_s \quad \text{and} \quad \bar{M}_t = \int_{T-t}^T \nabla w_\eta(Y_s) \cdot d\bar{W}_s,$$

we deduce from adding (39) and (40) that

$$\begin{aligned} 0 &= \int_0^t Lw_\eta(X_s) ds + \int_{T-t}^T Lw_\eta(Y_s) ds + \sqrt{2\beta^{-1}} (M_t + \bar{M}_t) \\ &= \int_0^t Lw_\eta(X_s) ds + \int_0^t Lw_\eta(Y_{T-s}) ds + \sqrt{2\beta^{-1}} (M_t + \bar{M}_t) \\ &= 2 \int_0^t Lw_\eta(X_s) ds + \sqrt{2\beta^{-1}} (M_t + \bar{M}_t). \end{aligned}$$

Hence

$$4\mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t Lw_\eta(X_s) ds \right|^2 \right] \leq 4\beta^{-1} \left(\mathbb{E} \left[\sup_{0 \leq t \leq T} |M_t|^2 \right] + \mathbb{E} \left[\sup_{0 \leq t \leq T} |\bar{M}_t|^2 \right] \right).$$

The random process $(M_t)_{0 \leq t \leq T}$ is a martingale since $\mathbb{E} \int_0^T |\nabla w_\eta|^2(X_s) ds = T \int_{\mathbb{R}^n} |\nabla w_\eta|^2 \psi < \infty$ in view of (37) and the fact that, for any time $s \in [0, T]$, X_s is distributed according to μ . Using Doob inequality on the martingales $(M_t)_{0 \leq t \leq T}$ and $(\overline{M}_t)_{0 \leq t \leq T}$, which reads

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |M_t|^2 \right] \leq 4 \mathbb{E} [|M_T|^2]$$

and likewise for \overline{M}_t , we obtain

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t Lw_\eta(X_s) ds \right|^2 \right] &\leq 8\beta^{-1} \mathbb{E} \left[\int_0^T |\nabla w_\eta(X_s)|^2 ds \right] \\ &= 8\beta^{-1} T \int_{\mathbb{R}^n} |\nabla w_\eta|^2 \psi \\ &\leq 8\beta T \|\Phi\|_{L^2(\psi)}^2, \end{aligned} \tag{41}$$

where we have used (37) in the last line.

In view of (36), we now write, for any $\nu > 0$, that

$$\left| \int_0^t \nabla^* \Phi(X_s) ds \right|^2 \leq (1 + \nu) \left| \int_0^t Lw_\eta(X_s) ds \right|^2 + \left(1 + \frac{1}{\nu} \right) \left| \int_0^t \eta w_\eta(X_s) ds \right|^2,$$

thus

$$\begin{aligned} &\mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t \nabla^* \Phi(X_s) ds \right|^2 \right] \\ &\leq (1 + \nu) \mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t Lw_\eta(X_s) ds \right|^2 \right] + \left(1 + \frac{1}{\nu} \right) \mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t \eta w_\eta(X_s) ds \right|^2 \right] \\ &\leq (1 + \nu) 8\beta T \|\Phi\|_{L^2(\psi)}^2 + \left(1 + \frac{1}{\nu} \right) \eta^2 T \mathbb{E} \left[\int_0^T w_\eta^2(X_s) ds \right] \\ &\leq (1 + \nu) 8\beta T \|\Phi\|_{L^2(\psi)}^2 + \left(1 + \frac{1}{\nu} \right) \eta^2 T^2 \|w_\eta\|_{L^2(\psi)}^2. \end{aligned}$$

We now pass to the limit $\eta \rightarrow 0$ using (38). We get that, for any $\nu > 0$,

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t \nabla^* \Phi(X_s) ds \right|^2 \right] \leq (1 + \nu) 8\beta T \|\Phi\|_{L^2(\psi)}^2,$$

which implies (35). This concludes the proof of Lemma 10. \square

4.3 Estimate on the term $\int_0^t L^{X_s^1} u(X_s) ds$

We are now in position to bound the last term in (24).

Proposition 11. *Let X_t be the solution to (1), with its initial condition distributed according to the equilibrium measure μ (see assumption (17)). We assume that (10) and (12) hold. Let u be defined as the solution to the Poisson equation (31). Then, we have*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t L^{X_s^1} u(X_s) ds \right|^2 \right] \leq 8T\beta \frac{\kappa^2}{\rho^2}. \quad (42)$$

Proof. Let us introduce $\Phi = (0, \widehat{\nabla} u) : \mathbb{R}^n \rightarrow \mathbb{R}^n$. We have

$$\begin{aligned} L^\xi u &= -\widehat{\nabla} V \cdot \widehat{\nabla} u + \beta^{-1} \widehat{\Delta} u \\ &= \beta^{-1} (-\beta \widehat{\nabla} V \cdot \widehat{\nabla} u + \widehat{\Delta} u) \\ &= -\beta^{-1} \nabla^\star \Phi. \end{aligned}$$

Since $u \in \mathcal{C}^\infty$, we have $\Phi \in (\mathcal{C}^\infty)^n$. Furthermore, in view of (33), we have $\Phi \in (L^2(\psi))^n$. We are thus in position to use (35) and get that

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t L^{X_s^1} u(X_s) ds \right|^2 \right] \leq 8T\beta^{-1} \int_{\mathbb{R}^n} |\widehat{\nabla} u|^2 \psi.$$

The estimate (42) is then obtained as a direct consequence of the estimate (33) on the solution u to the Poisson problem. \square

5 A Gronwall-type result

In this section, we state a general Gronwall-type result that will be crucial to prove Proposition 3. This section explains the role of the one-sided Lipschitz assumption (13) as well as the integrability condition (15).

Lemma 12. *Consider a smooth function $b : \mathbb{R} \rightarrow \mathbb{R}$ satisfying the one-sided Lipschitz assumption (13) with constant L_b and the integrability condition (15) with constant C_α . Let $X_t \in \mathbb{R}$ and $Y_t \in \mathbb{R}$ be the solutions to*

$$dX_t = -b(X_t)dt + \sigma dB_t, \quad (43)$$

$$dY_t = -b(Y_t)dt + \sigma dB_t + f_t dt, \quad (44)$$

for some time integrable stochastic process f_t and some positive constant σ , where B_t is a one-dimensional Brownian motion. We assume that there exists a probability measure $\varphi(x) dx$ on \mathbb{R} which is invariant both for the dynamics (43) and (44) and that $X_0 = Y_0$ are distributed according to that measure.

Consider a fixed time interval $[0, t]$. We have

$$\mathbb{E} \left[\sup_{s \in [0, t]} |X_s - Y_s| \right] \leq C \sqrt{\mathbb{E} \left[\sup_{s \in [0, t]} |e_s|^2 \right]} \quad (45)$$

where $e_s = \int_0^s f_\tau d\tau$ and C is a constant only depending on t , C_α and L_b .

Let us comment on this result. First, using the one-sided assumption (13), one can check that $d|X_t - Y_t| \leq (L_b|X_t - Y_t| + |f_t|) dt$, which yields, by the Gronwall lemma, an estimate of the difference $(|X_s - Y_s|)_{s \in [0, t]}$ in terms of $(|f_s|)_{s \in [0, t]}$. Here, we only assume a control on $\left(\left| \int_0^s f_\tau d\tau \right| \right)_{s \in [0, t]}$. This is why the standard approach does not apply, and why we need Assumption (15) in addition to the one-sided Lipschitz assumption (13).

Second, if b is assumed to be globally Lipschitz with constant L_b (instead of one-sided Lipschitz), one can check that $|X_t - Y_t| \leq L_b \int_0^t |X_s - Y_s| ds + |e_t|$ and, again by Gronwall lemma, $\sup_{s \in [0, t]} |X_s - Y_s| \leq e^{L_b t} \sup_{s \in [0, t]} |e_s|$. Here, we do not assume b Lipschitz but only one-sided Lipschitz, and we were not able to prove such a pathwise inequality under the only assumption (13). Actually, we can prove a similar inequality but only in expectation, see (45), and under the additional assumption (15).

Concerning the dependency of the constant C in (45) on the time t , one can check from the proof that $C = 1 + \tilde{C}t \exp(L_b t)$ where \tilde{C} is a constant only depending on C_α and L_b . Notice that if b is assumed to be increasing, one can take $L_b = 0$, so that $C = 1 + \tilde{C}t$.

Before proving Lemma 12, let us comment on Assumption (15), that reads, we recall,

$$C_\alpha(\beta) = \int_{\mathbb{R}} \left(\sup_{s \in [-|x|, |x|]} |b'(s)| \right)^2 \varphi(x) dx < \infty.$$

Remark 13. *It is not enough to assume*

$$\int_{\mathbb{R}} |b'(x)|^2 \varphi(x) dx < \infty \quad (46)$$

for (15) to hold. Consider indeed the case of $\varphi(x) = (1 + |x|^2)^{-1}$ and $b' = 0$ on \mathbb{R} , except in $[n, n + 1/n^2]$, where $|b'| = n$, for any $n \in \mathbb{N}^*$. Then

$$\int_{\mathbb{R}} |b'(x)|^2 \varphi(x) dx \leq \sum_{n \geq 1} n^2 (1 + |n|^2)^{-1} \frac{1}{n^2} < \infty,$$

and (46) holds. But, for any x such that $n < |x| < n + 1$, $\sup_{s \in [-|x|, |x|]} |b'(s)| = n$,

thus

$$\int_{\mathbb{R}} \left(\sup_{s \in [-|x|, |x|]} |b'(s)| \right)^2 \varphi(x) dx \geq 2 \sum_{n \geq 0} n^2 \frac{1}{1 + |n + 1|^2} = \infty$$

and (15) does not hold.

Remark 14. *Rather than assuming (15), another possibility is to assume that b' satisfies (46) and that there exists C such that, for any x and y ,*

$$\sup_{\theta \in (x, y)} |b'(\theta)| \leq C \left(|b'(x)| + |b'(y)| \right).$$

Then (45) again holds.

Proof of Lemma 12. As an obvious consequence of the assumptions, we see that, at any time t , X_t and Y_t share the same probability law $\varphi(x) dx$, independent of t . Let $Z_t = Y_t - e_t$. We infer from (43) and (44) that

$$\begin{aligned} d|X_t - Z_t| &= (b(Z_t + e_t) - b(X_t)) \operatorname{sgn}(X_t - Z_t) dt \\ &= (1_{|X_t - Z_t| < |e_t|} + 1_{|X_t - Z_t| \geq |e_t|}) (b(Z_t + e_t) - b(X_t)) \operatorname{sgn}(X_t - Z_t) dt. \end{aligned} \quad (47)$$

To control the second term in the above right-hand side (corresponding to the case $|X_t - Z_t| \geq |e_t|$), we argue as follows:

- If $X_t \geq Z_t$, then it means that $X_t - Z_t \geq |e_t| \geq e_t$, hence $X_t \geq Z_t + e_t$, hence, using (13), we have $b(Z_t + e_t) - b(X_t) \leq L_b(X_t - Z_t - e_t)$.

Therefore,

$$\begin{aligned}
& 1_{X_t \geq Z_t} 1_{|X_t - Z_t| \geq |e_t|} (b(Z_t + e_t) - b(X_t)) \operatorname{sgn}(X_t - Z_t) \\
&= (b(Z_t + e_t) - b(X_t)) 1_{X_t - Z_t \geq |e_t|} \\
&\leq L_b(X_t - Z_t - e_t) 1_{X_t - Z_t \geq |e_t|} \\
&\leq L_b(|X_t - Z_t| + |e_t|).
\end{aligned} \tag{48}$$

- If $X_t \leq Z_t$, then it means that $X_t - Z_t \leq -|e_t| \leq e_t$, hence $X_t \leq Z_t + e_t$, hence, using (13), we get $b(Z_t + e_t) - b(X_t) \geq L_b(X_t - Z_t - e_t)$. Therefore,

$$\begin{aligned}
& 1_{X_t \leq Z_t} 1_{|X_t - Z_t| \geq |e_t|} (b(Z_t + e_t) - b(X_t)) \operatorname{sgn}(X_t - Z_t) \\
&= -(b(Z_t + e_t) - b(X_t)) 1_{X_t - Z_t \leq -|e_t|} \\
&\leq -L_b(X_t - Z_t - e_t) 1_{X_t - Z_t \leq -|e_t|} \\
&\leq L_b(|X_t - Z_t| + |e_t|).
\end{aligned} \tag{49}$$

Collecting (47), (48) and (49), we have

$$d|X_t - Z_t| \leq 1_{|X_t - Z_t| < |e_t|} |b(X_t) - b(Z_t + e_t)| dt + L_b(|X_t - Z_t| + |e_t|) dt. \tag{50}$$

To proceed, we write

$$|b(X_t) - b(Z_t + e_t)| = |b(X_t) - b(Y_t)| = |X_t - Y_t| |b'(\theta_t)|$$

for some $\theta_t \in (X_t, Y_t)$. Using the function α defined by (14), we obtain

$$|b(X_t) - b(Z_t + e_t)| \leq |X_t - Y_t| (\alpha(|X_t|) + \alpha(|Y_t|)).$$

Inserting this relation in (50), we obtain

$$\begin{aligned}
& d|X_t - Z_t| \\
&\leq 1_{|X_t - Z_t| < |e_t|} |X_t - Y_t| (\alpha(|X_t|) + \alpha(|Y_t|)) dt + L_b(|X_t - Z_t| + |e_t|) dt \\
&= 1_{|X_t - Z_t| < |e_t|} |X_t - Z_t - e_t| (\alpha(|X_t|) + \alpha(|Y_t|)) dt + L_b(|X_t - Z_t| + |e_t|) dt \\
&\leq 2 (\alpha(|X_t|) + \alpha(|Y_t|) + L_b/2) |e_t| dt + L_b |X_t - Z_t| dt.
\end{aligned}$$

We thus obtain

$$d \left[\exp(-L_b t) |X_t - Z_t| \right] \leq 2 \exp(-L_b t) (\alpha(|X_t|) + \alpha(|Y_t|) + L_b/2) |e_t| dt.$$

We now integrate in time, and use the fact that $X_0 = Z_0$:

$$\begin{aligned} \exp(-L_b t) |X_t - Z_t| &\leq 2 \int_0^t \exp(-L_b s) (\alpha(|X_s|) + \alpha(|Y_s|) + L_b/2) |e_s| ds \\ &\leq 2 \left(\sup_{s \in [0, t]} |e_s| \right) \int_0^t (\alpha(|X_s|) + \alpha(|Y_s|) + L_b/2) ds. \end{aligned}$$

We deduce that

$$\sup_{s \in [0, t]} |X_s - Z_s| \leq 2 \exp(L_b t) \left(\sup_{s \in [0, t]} |e_s| \right) \int_0^t (\alpha(|X_s|) + \alpha(|Y_s|) + L_b/2) ds, \quad (51)$$

thus

$$\begin{aligned} &\mathbb{E} \left[\sup_{s \in [0, t]} |X_s - Z_s| \right] \\ &\leq 2 \exp(L_b t) \sqrt{\mathbb{E} \left[\sup_{s \in [0, t]} |e_s|^2 \right]} \sqrt{\mathbb{E} \left[\int_0^t (\alpha(|X_s|) + \alpha(|Y_s|) + L_b/2) ds \right]^2} \\ &\leq 2 \exp(L_b t) \eta \sqrt{t} \sqrt{\mathbb{E} \left[\int_0^t \left(\alpha(|X_s|) + \alpha(|Y_s|) + L_b/2 \right)^2 ds \right]} \\ &\leq 2 \exp(L_b t) \eta \sqrt{3t} \sqrt{\mathbb{E} \left[\int_0^t \left[\left(\alpha(|X_s|) \right)^2 + \left(\alpha(|Y_s|) \right)^2 + \frac{L_b^2}{4} \right] ds \right]} \end{aligned}$$

where $\eta = \sqrt{\mathbb{E} [\sup_{s \in [0, t]} |e_s|^2]}$. Since the law of X_s and Y_s is the same and does not depend on time, $\mathbb{E} [\alpha(|X_s|)^2] = \mathbb{E} [\alpha(|Y_s|)^2] = \mathbb{E} [\alpha(|X_0|)^2]$ so that

$$\mathbb{E} \left[\sup_{s \in [0, t]} |X_s - Z_s| \right] \leq \tilde{C} t \exp(L_b t) \eta, \quad (52)$$

where \tilde{C} only depends on C_α and L_b . We eventually write that

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [0, t]} |X_s - Y_s| \right] &= \mathbb{E} \left[\sup_{s \in [0, t]} |X_s - Z_s - e_s| \right] \\ &\leq \mathbb{E} \left[\sup_{s \in [0, t]} |X_s - Z_s| \right] + \mathbb{E} \left[\sup_{s \in [0, t]} |e_s| \right]. \end{aligned}$$

From (52) and the above bound, we deduce the claimed bound (45). \square

Referring to Remark 4, if we assume (16) for some $1 \leq p < 2$ rather than (15), the above proof can be modified to prove

$$\mathbb{E} \left[\sup_{s \in [0, t]} |X_s - Y_s|^p \right] \leq C \left(\mathbb{E} \left[\sup_{s \in [0, t]} |e_s|^2 \right] \right)^{p/2} \quad (53)$$

instead of (45). Indeed, from (51), we deduce that

$$\sup_{s \in [0, t]} |X_s - Z_s|^p \leq 2^p \exp(pL_b t) \left(\sup_{s \in [0, t]} |e_s|^p \right) \left(\int_0^t (\alpha(|X_s|) + \alpha(|Y_s|) + L_b/2) ds \right)^p,$$

thus, using Hölder inequality with exponents $2/p$ and $2/(2-p)$, we get

$$\begin{aligned} & \mathbb{E} \left[\sup_{s \in [0, t]} |X_s - Z_s|^p \right] \\ & \leq C_1 \left(\mathbb{E} \left[\sup_{s \in [0, t]} |e_s|^2 \right] \right)^{p/2} \left(\mathbb{E} \left[\int_0^t (\alpha(|X_s|) + \alpha(|Y_s|) + L_b/2) ds \right]^{2p/(2-p)} \right)^{1-p/2} \\ & \leq C_2 \eta^p \left(\mathbb{E} \left[\int_0^t \left[(\alpha(|X_s|))^{2p/(2-p)} + (\alpha(|Y_s|))^{2p/(2-p)} + 1 \right] ds \right] \right)^{1-p/2}, \end{aligned}$$

where, as above, $\eta = \sqrt{\mathbb{E} [\sup_{s \in [0, t]} |e_s|^2]}$ and C_1 and C_2 are constants only depending on p , t and L_b . Since the law of X_s and Y_s is the same and does not depend on time,

$$\mathbb{E} [\alpha(|X_s|)^{2p/(2-p)}] = \mathbb{E} [\alpha(|Y_s|)^{2p/(2-p)}] = \mathbb{E} [\alpha(|X_0|)^{2p/(2-p)}] = C_{\alpha, p} < \infty$$

in view of (16). We hence have $\mathbb{E} \left[\sup_{s \in [0, t]} |X_s - Z_s|^p \right] \leq C \eta^p$, where C only depends on p , t , $C_{\alpha, p}$ and L_b . From this estimate and the fact that $X_s - Y_s = X_s - Z_s - e_s$, we deduce (53).

6 Proof of Proposition 3

In this section, we complete the proof of the error estimate (18) between the effective dynamics and the original dynamics, by combining the estimate on

the last term in (24) obtained in Section 4 (namely (42)) together with the Gronwall-type argument of Section 5 (i.e. estimate (45)).

Recall that the exact dynamics satisfies (see (22))

$$X_t^1 - X_0^1 = - \int_0^t b(X_s^1) ds + \sqrt{2\beta^{-1}} W_t^1 + \int_0^t L^{X_s^1} u(X_s) ds$$

whereas the effective dynamics satisfies (see (23))

$$\xi_t - X_0^1 = - \int_0^t b(\xi_s) ds + \sqrt{2\beta^{-1}} W_t^1.$$

We set

$$f_s = L^{X_s^1} u(X_s), \quad e_t = \int_0^t L^{X_s^1} u(X_s) ds,$$

and we are thus in the setting of Lemma 12, where ξ_t satisfies a dynamics of the form (43) while X_t^1 satisfies a dynamics of the form (44).

In view of the assumptions (13), (15) and (17), we see that the assumptions of Lemma 12 are satisfied. The bound (45) thus yields that

$$\mathbb{E} \left[\sup_{s \in [0, T]} |\xi_s - X_s^1| \right] \leq C\eta, \quad (54)$$

where

$$\eta^2 = \mathbb{E} \left[\sup_{0 \leq t \leq T} |e_t|^2 \right] = \mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t L^{X_s^1} u(X_s) ds \right|^2 \right] \leq 8T\beta \frac{\kappa^2}{\rho^2},$$

the last inequality being the main output of Section 4 (see (42) in Proposition 11). Inserting the above estimate in (54), we deduce the claimed bound (18). This concludes the proof of Proposition 3.

Remark 15. *Following Remark 4, we now replace the assumption (15) by the stronger assumption (16) for some $1 \leq p < 2$. Following the same lines as in the proof of Proposition 3, and using (53) instead of (45), we obtain*

$$\mathbb{E} \left[\sup_{s \in [0, T]} |\xi_s - X_s^1|^p \right] \leq C\eta^p, \quad (55)$$

where, in view of (42), we have $\eta^2 \leq 8T\beta \frac{\kappa^2}{\rho^2}$. Inserting this estimate in (55), we deduce (19).

7 Application: a quantitative averaging result

In this section, as an application of the techniques presented above, we show how to obtain a quantitative result on the error introduced by an averaging principle. We would like to stress that our result holds without assuming that the effective drift function is Lipschitz, which is to the best of our knowledge the assumption made in similar quantitative results that have been previously obtained in the literature.

Let us consider the stochastic differential equation

$$\begin{cases} dX_t^{\varepsilon,1} = -\partial_1 V(X_t^\varepsilon) dt + \sqrt{2\beta^{-1}} dW_t^1 \\ dX_t^{\varepsilon,i} = -\frac{1}{\varepsilon} \partial_i V(X_t^\varepsilon) dt + \sqrt{\frac{2\beta^{-1}}{\varepsilon}} dW_t^i \end{cases} \quad \text{for } i = 2, \dots, n \quad (56)$$

where $X_t^{\varepsilon,i}$ denotes the i -th component of the vector $X_t^\varepsilon \in \mathbb{R}^n$ and ε is a positive constant. The initial condition $X_0^\varepsilon = X_0$ is assumed to be independent of ε for simplicity. We note that $d\mu = \psi(x) dx$ is again the invariant measure of (56), and that it is independent of ε .

By the averaging principle (see for example [15]), it is expected that, in the limit $\varepsilon \rightarrow 0$, the process $(X_t^{\varepsilon,1})_{t \geq 0}$ converges to the process $(\xi_t)_{t \geq 0}$ satisfying

$$d\xi_t = -b(\xi_t) dt + \sqrt{2\beta^{-1}} dW_t^1 \quad (57)$$

with initial condition $\xi_0 = X_0^1$, and where b is defined by (3). Using the techniques presented above, we are able to prove the following convergence result:

Proposition 16. *Assume that (10), (12), (13) and (15) hold, and that the system starts at equilibrium:*

$$X_0 \sim \mu. \quad (58)$$

Consider $(X_t)_{0 \leq t \leq T}$ solution to (56) and $(\xi_t)_{0 \leq t \leq T}$ solution to (57) over a bounded time interval $[0, T]$. Then, there exists a constant C , that is independent of ε , ρ and κ , and that only depends on T , $C_\alpha(\beta)$ and L_b , such that

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^{\varepsilon,1} - \xi_t| \right) \leq C \sqrt{\beta \varepsilon} \frac{\kappa}{\rho}. \quad (59)$$

Proof. The proof consists in exactly following the same arguments as for the proof of Proposition 3, keeping track of the dependency of the constants on ε . Let us emphasize the modifications in the various steps of the proof.

The equation (24) still holds in our context:

$$X_t^{\varepsilon,1} - \xi_t = \int_0^t (b(\xi_s) - b(X_s^{\varepsilon,1})) ds + \int_0^t L^{X_s^{\varepsilon,1}} u(X_s^\varepsilon) ds, \quad (60)$$

where, as before, u is the solution to the Poisson problem (21) using the same definition (9) for the operator L^ξ . In particular, u does not depend on ε . By using the results of Sections 3 and 4.1, we have (see Lemma 9) that

$$\int_{\mathbb{R}^n} |\widehat{\nabla} u|^2 \psi \leq \beta^2 \frac{\kappa^2}{\rho^2}. \quad (61)$$

The infinitesimal generator of the process $(X_t^\varepsilon)_{t \geq 0}$ is L_ε defined by: for any smooth function $v : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$L_\varepsilon v = -\partial_1 V \partial_1 v + \beta^{-1} \partial_{11} v + \frac{1}{\varepsilon} L^\xi v,$$

to be compared with (8).

The main modification in the proof of Proposition 3 is in Lemma 10. Now, the estimate (35) is the following. For any function $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ in $(\mathcal{C}^\infty \cap L^2(\psi))^n$, we have

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t \nabla^* \Phi(X_s) ds \right|^2 \right] \leq 8T\beta \left(\|\Phi_1\|_{L^2(\psi)}^2 + \varepsilon \sum_{i=2}^n \|\Phi_i\|_{L^2(\psi)}^2 \right) \quad (62)$$

where Φ_i denotes the i -th component of Φ and, as above, $\nabla^* \Phi = \beta \nabla V \cdot \Phi - \operatorname{div}(\Phi)$. Let us give a few details on how (62) is obtained, mimicking the proof of Lemma 10. The equation $\eta w_{\varepsilon,\eta} - L_\varepsilon w_{\varepsilon,\eta} = -\nabla^* \Phi$ (compare with (36)) has a unique solution $w_{\varepsilon,\eta} \in H^1(\psi)$. The associated variational formulation is: for any $v \in H^1(\psi)$,

$$\eta \int_{\mathbb{R}^n} w_{\varepsilon,\eta} v \psi + \beta^{-1} \int_{\mathbb{R}^n} \left(\partial_1 w_{\varepsilon,\eta} \partial_1 v + \frac{1}{\varepsilon} \sum_{i=2}^n \partial_i w_{\varepsilon,\eta} \partial_i v \right) \psi = - \int_{\mathbb{R}^n} \Phi \cdot \nabla v \psi.$$

Taking $v = w_{\varepsilon,\eta}$, we obtain the a priori estimate (compare with (37)–(38))

$$\|\partial_1 w_{\varepsilon,\eta}\|_{L^2(\psi)}^2 + \frac{1}{\varepsilon} \sum_{i=2}^n \|\partial_i w_{\varepsilon,\eta}\|_{L^2(\psi)}^2 \leq \beta^2 \left(\|\Phi_1\|_{L^2(\psi)}^2 + \varepsilon \sum_{i=2}^n \|\Phi_i\|_{L^2(\psi)}^2 \right) \quad (63)$$

and

$$\lim_{\eta \rightarrow 0} \eta \|w_{\varepsilon, \eta}\|_{L^2(\psi)} = 0. \quad (64)$$

Following the Step 2 of the proof of Lemma 10, we write, by Itô's calculus, that

$$\begin{aligned} w_{\varepsilon, \eta}(X_t^\varepsilon) - w_{\varepsilon, \eta}(X_0) &= \int_0^t L_\varepsilon w_{\varepsilon, \eta}(X_s^\varepsilon) ds \\ &\quad + \sqrt{2\beta^{-1}} \int_0^t \left(\partial_1 w_{\varepsilon, \eta}(X_s^\varepsilon) dW_s^1 + \frac{1}{\sqrt{\varepsilon}} \sum_{i=2}^n \partial_i w_{\varepsilon, \eta}(X_s^\varepsilon) dW_s^i \right), \end{aligned}$$

which is to be compared with (39). The process $(X_s^\varepsilon)_{0 \leq s \leq T}$ is still reversible, and thus, one obtains, using (63), that

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t L_\varepsilon w_{\varepsilon, \eta}(X_s^\varepsilon) ds \right|^2 \right] &\leq 8\beta^{-1} T \int_{\mathbb{R}^n} \left((\partial_1 w_\varepsilon)^2 + \frac{1}{\varepsilon} \sum_{i=2}^n (\partial_i w_\varepsilon)^2 \right) \psi \\ &\leq 8T\beta \left(\|\Phi_1\|_{L^2(\psi)}^2 + \varepsilon \sum_{i=2}^n \|\Phi_i\|_{L^2(\psi)}^2 \right), \end{aligned}$$

to be compared with (41). Using (64), we obtain (62) as in the proof of Lemma 10.

Choosing $\Phi = (0, \widehat{\nabla} u)$ in (62) (as in the proof of Proposition 11), one obtains, using (61),

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t L^{X_s^{\varepsilon, 1}} u(X_s^\varepsilon) ds \right|^2 \right] \leq 8T\beta\varepsilon \frac{\kappa^2}{\rho^2},$$

which is to be compared with (42). This estimates gives the magnitude of last term in (60). The end of the proof follows exactly the same lines as for Proposition 3. \square

Let us make three comments on the previous result. First, the stationarity assumption (58) can be weakened in a similar way as in Corollary 6 above by using the same conditioning argument.

Second, it is easy to generalize the previous result to the situation where each component of $(X_t)_{t \geq 0}$ is scaled in time with a parameter ε_i : for $i = 2, \dots, n$,

$$dX_t^i = -\frac{1}{\varepsilon_i} \partial_i V(X_t) dt + \sqrt{\frac{2\beta^{-1}}{\varepsilon_i}} dW_t^i$$

while we keep, for the first component,

$$dX_t^1 = -\partial_1 V(X_t) dt + \sqrt{2\beta^{-1}} dW_t^1.$$

In this case, one obtains a similar estimate as (59), which ε being replaced by $\max(\varepsilon_2, \dots, \varepsilon_n)$.

Third, we notice that using the simple approach of Section 3.2 on the stochastic differential equation (56), one obtains the estimate (29) with an upper bound independent of ε (recall indeed that b , and thus f defined by (20), are independent of ε). This shows the interest of the approach developed in Sections 4 and 5.

Acknowledgments

S. Olla acknowledges support by the ANR LSD. The work of F. Legoll and T. Lelièvre is supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement number 614492. T. Lelièvre would like to thank Dirk Blömker (Universität Augsburg) for useful discussions on a preliminary version of this work.

References

- [1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*. Société Mathématique de France, 2000. In French.
- [2] W. E and E. Vanden-Eijnden. Metastability, conformation dynamics, and transition pathways in complex systems. In *Multiscale modelling and simulation*, volume 39 of *Lect. Notes Comput. Sci. Eng.*, pages 35–68. Springer, Berlin, 2004.
- [3] D. Givon, R. Kupferman, and A. Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, 17(6):R55–R127, 2004.
- [4] N. Grunewald, F. Otto, C. Villani, and M.G. Westdickenberg. A two-scale approach to logarithmic Sobolev inequalities and the hydrody-

- namical limit. *Ann. Inst. H. Poincaré Probab. Statist.*, 45(2):302–351, 2009.
- [5] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probab. Th. Rel. Fields*, 71:501–516, 1986.
 - [6] P. Hänggi, P. Talkner, and M. Borkovec. Reaction-rate theory: fifty years after Kramers. *Reviews of Modern Physics*, 62(2):251–342, 1990.
 - [7] T. Komorowski, C. Landim, and S. Olla. *Fluctuations in Markov Processes, Time Symmetry and Martingale Approximation*, volume 345 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Berlin, New York, 2012.
 - [8] F. Legoll and T. Lelièvre. Effective dynamics using conditional expectations. *Nonlinearity*, 23:2131–2163, 2010.
 - [9] F. Legoll and T. Lelièvre. Some remarks on free energy and coarse-graining. In B. Engquist, O. Runborg, and R. Tsai, editors, *Multiscale Modeling and Simulation in Science*, volume 82 of *Lecture Notes in Computational Science and Engineering*, pages 279–329. Springer, 2012.
 - [10] T. Lelièvre. A general two-scale criteria for logarithmic Sobolev inequalities. *J. Funct. Anal.*, 256(7):2211–2221, 2009.
 - [11] T. Lelièvre, M. Rousset, and G. Stoltz. Long-time convergence of an adaptive biasing force method. *Nonlinearity*, 21:1155–1181, 2008.
 - [12] T. Lelièvre, M. Rousset, and G. Stoltz. *Free energy computations: A mathematical perspective*. Imperial College Press, 2010.
 - [13] T.J. Lyons and T.S. Zhang. Decomposition of Dirichlet processes and its application. *The Annals of Probability*, 22(1):494–524, 1994.
 - [14] L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti. String method in collective variables: minimum free energy paths and isocommittor surfaces. *J. Chem. Phys.*, 125:024106, 2006.
 - [15] G.A. Pavliotis and A.M. Stuart. *Multiscale methods: averaging and homogenization*. Springer, 2007.